Accident Analysis and Prevention 135 (2020) 105390

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Identifying the causes of road traffic collisions: Using police officers' expertise to improve the reporting of contributory factors data

Jonathan J. Rolison*

Department of Psychology, University of Essex, UK

ARTICLE INFO

Keywords:
Road safety
Crash causation
Contributory factors
Policymaking
Crash reporting

ABSTRACT

Worldwide, road crashes are a major course of death and serious injury. Police reports provide a rich source of data on the proximal causes (e.g., impairment by alcohol, failure to look properly) of road traffic collisions. Yet, road safety research has raised concerns about the quality and reliability of police reported data. In the UK crash report form, contributory factors are categorised (e.g., vehicle defects, driver error or reaction) to aid police officers in identifying appropriate factors. However, discord between the classification of contributory factors in crash reports and police officers' own categorical perceptions may lead to misunderstanding, and in turn, misreporting of contributory factors. The current investigation recruited 162 police officers to report their perceptions of the relations among contributory factors in the UK crash report form. Hierarchical clustering analysis was used to identify an optimal category structure based on police officers' perceptions. The clustering analysis identified a classification system with seven or eleven categories of contributory factors, maximising the internal coherence of categories and minimising discord with police officers' perceptions. The findings also yield new insights into police officers' perceptions of crash causation and demonstrate how statistical techniques can be used to inform the design of road traffic collision report forms.

1. Introduction

Road crashes are a major cause of death worldwide, accounting for more than 1.2 million deaths each year and many more non-fatal injuries (World Health Organization, 2015). Driver-related factors, namely driver actions or behaviour, contribute to most road traffic collisions and are the dominant cause of the majority of crashes (Evans, 1996). While there are multiple routes to improving road safety, such as by improving the road environment (Weijermars and Wegman, 2011), many driver-related factors are preventable (e.g., temporary distraction, exceeding the speed limit; Petridou and Moustaki, 2000; Rolison et al., 2018), implying that on the basis of reliable data about the factors that contribute to crashes road safety policies and initiatives could further be informed to improve public safety (Elder et al., 2004; Shope, 2007). Police reports provide an important source of data on the factors that contribute to road traffic collisions (e.g.,UK Department for Transport (DfT, 2014). However, road safety research has raised concerns about the quality and reliability of police reported data (Couto et al., 2016; Watson et al., 2015), especially regarding the reporting of contributory factors to crashes (Imprialou and Quddus, 2019; Montella, 2011; Rolison et al., 2018), indicating that road traffic collision reports may provide a misleading picture about crash causation. The current research investigated police officers' perceptions of existing methods for reporting contributory factors with a view to improving the quality and reliability of police reported data.

A wealth of research investigating the factors that contribute to road traffic collisions is based on police reported data (e.g., Gonzales et al., 2005; Lam, 2003; Langford and Koppel, 2006; McGwin and Brown, 1999). Police reports provide a rich source of nationally representative crash causation data, in comparison with small scale in-depth collision investigations (Beanland et al., 2013b, 2013a; Larsen, 2004). Police officers who attend a road traffic collision provide a subjective report of the factors that they believe contributed to the crash. For example, in the United Kingdom (UK), police officers who attend an incident provide a subjective report of the factors that they believe contributed to the crash. To do so, officers select among various possible contributors, categorised as road environment (e.g., animal or object in the carriageway), vehicle defects (e.g., defective steering or suspension), injudicious action (e.g., following too close), error or reaction (e.g., poor turn or manoeuvre), impairment or distraction (e.g., fatigue), behaviour or inexperience (e.g., uncertain, nervous or panic), and vision affected (e.g., dazzling sun; UK DfT, 2018a; see Table 1 for a full list of factors). The categories used to classify contributory factors in the UK crash report are intended as a classification system to aid officers in

* Corresponding author at: Department of Psychology, University of Essex, Essex, CO4 3SG, UK.
E-mail address: jrolison@essex.ac.uk.

https://doi.org/10.1016/j.aap.2019.105390
Received 6 August 2019; Received in revised form 26 November 2019; Accepted 30 November 2019

Table 1
Contributory factors in the UK road crash report form.

	Seven-cluster solution	Eleven-cluster solution
Road environment contributed		
Poor or defective road surface	7	11
Deposit on road (e.g., oil, mud, chippings)	7	11
Slippery road (due to weather)	7	11
Inadequate or masked signs or road markings	7	11
Defective traffic signals	7	11
Traffic calming (e.g., speed cushions, road humps, chicanes)	7	11
Temporary road layout (e.g., contraflow)	7	11
Road layout (e.g., bend, hill, narrow carriageway)	7	11
Animal or object in carriageway	7	11
Sunken, raised or slippery inspection cover	7	11
Vehicle defects		
Tyres illegal, defective or under-inflated	3	5
Defective lights or indicators	3	5
Defective brakes	3	5
Defective steering or suspension	3	5
Defective or missing mirrors	3	5
Overloaded or poorly loaded vehicle or trailer	3	5
Injudicious action		
Disobeyed automatic traffic signal	2	4
Disobeyed Give Way or Stop sign or markings	2	4
Disobeyed double white lines	2	4
Disobeyed pedestrian crossing facility	2	4
Illegal turn or direction of travel	2	4
Exceeding speed limit	2	4
Travelling too fast for conditions	2	3
Following too close	2	3
Vehicle travelling along pavement	2	4
Cyclist entering road from pavement	5	9
Driver error or reaction		
Junction overshoot	2	2
Junction restart (moving off at junction)	2	2
Poor turn or manoeuvre	2	2
Failed to signal or misleading signal	2	2
Failed to look properly	2	3
Failed to judge other person's path or speed	2	3
Too close to cyclist, horse or pedestrian	2	3
Sudden braking	2	2
Swerved	2	2
Loss of control	2	2
Impairment or distraction		
Impaired by alcohol	4	6
Impaired by drugs (illicit or medicinal)	4	6
Fatigue	4	8
Uncorrected, defective eyesight	4	8
Illness or disability, mental or physical	4	8
Not displaying lights at night or in poor visibility	3	5
Rider wearing dark clothing	5	9
Driver using mobile phone	4	6
Distraction in vehicle	4	7
Distraction outside vehicle	4	7
Behavior or Inexperience		
Aggressive driving	2	4
Careless, reckless or in a hurry	2	4
Nervous, uncertain or panic	1	1
Driving too slow for conditions or slow vehicle (e.g. tractor)	1	1
Learner or inexperienced driver	1	1
Inexperience of driving on the left	1	1
Unfamiliar with model of vehicle	1	1
Vision Affected by		
Stationary or parked vehicle(s)	6	10
Vegetation	6	10
Road layout (e.g. bend, winding road, hill crest)	6	10
Buildings, road signs, street furniture	6	10
Dazzling headlights	6	10
Dazzling sun	6	10
Rain, sleet, snow or fog	6	10
Spray from other vehicles	6	10
Visor or windscreen dirty, scratched or frosted etc.	3	5
Vehicle blind spot	6	10

identifying appropriate factors. For example, factors in the 'injudicious action' category are intended to relate specifically to injudicious actions or behaviours that are distinguishable from other types of actions or behaviours, such as those related to impairment or distraction (Table 1). The classification system is also intended to distinguish types of road user behaviours (e.g., injudicious action) from other behaviours (e.g., impairment or distraction) and non-driver related factors (e.g., vehicle defects). In this latter sense, the classification of contributory factors into superordinate categories helps focus road safety priorities on broad categories of crash causation (UK Department for Transport (DfT, 2014; 2017). Similarly, in the United States, drivers involved in a road traffic collision have been interviewed about the events leading up to a crash (e.g., driver distraction, exceeding the speed limit) in combination with investigation of the incident scene (NHTSA, 2008).

However, road safety research has raised concerns regarding the reliability of police reported contributory factors data (Imprialou & Quddus, in press; Montella, 2011; Rolison et al., 2018). One concern is that police officers' perceptions about some of the typical factors involved in road traffic collisions are at odds with the frequency that contributory factors are reported in crash records. Rolison et al. (2018) found that some contributory factors were under-represented (e.g., driver distraction) or over-represented (e.g., failed to look properly) in crash records with respect to police officers' perceptions of their occurrence in crashes. While some discrepancies may result from under-reporting of factors that are hard to verify by police officers (e.g., mobile phone use; Rolison et al., 2018; Montella, 2011; Watson et al., 2015), discrepancies between crash records and perceptions of police officers who complete those records suggest possible misreporting of certain contributory factors due to incomplete reporting (Imprialou & Quddus, in press) or misuse of the crash report form (Broughton et al., 2010). The current investigation focusses on the adequacy of the fixed list of contributory factors in the UK crash report form that police officers select among to identify factors contributing to crashes. It is intended that by focussing on the UK practices, identifying potential bias and possible improvements will inform collision reporting practices internationally.

In the UK crash report form, the 'behaviour or inexperience' category contains various driver-related characteristics and behaviours, including some that refer explicitly to driver inexperience (e.g., 'learner or inexperienced driver') and others that do not refer to level of driver experience (e.g., 'careless, reckless, or in a hurry'; UK DfT, 2018a; Table 1). It may be unclear to the reporting police officer which factors in this category should be considered in relation to inexperience. For instance, for some police officers 'aggressive driving' may only be identified as a factor contributing to road crashes involving inexperienced drivers, which would lead to inconsistent reporting across police officers. Moreover, it may be unclear whether this category represents a unified set of related factors that might be prioritised by road safety organisations and which factors it contains could be combined to represent a coherent driver characteristic or behaviour (e.g., driver inexperience).

Research in psychology has revealed that how items are categorised influences people's perception of the items that categories contain. When categories are imposed, perceived differences between items in

the same category are minimised, whereas perceived differences between items in separate categories are accentuated (Goldstone, 1994; Goldstone and Hendrickson, 2010; Schusterman et al., 2000). This tendency, known as categorical perception, illustrates how the perception of items (e.g., contributory factors) is influenced by imposed categories (e.g., categories used to classify contributory factors). Perceptually dissimilar items can be perceived as more similar, or even equivalent, if classified in the same category (Schusterman et al., 2000). Thus, it is crucial that when categories are used to classify fixed lists of contributory factors in a crash report that the categories contain coherent sets of similar items, maximising dissimilarity between categories. Otherwise, items contained in diverse or eclectic categories are likely to be perceived by police officers as more similar or related due to their classification in the same category. Moreover, it is important that contributory factor categories map onto police officers' categorical perceptions of the factors. Discord between the imposed classification of contributory factors and police officers' own categorical perceptions could lead to erroneous or inconsistent reporting due to misunderstanding of the meaning of individual factors.

In sum, it is important that when road traffic collision reports contain a fixed set of contributory factors that the factors are classified into coherent categories to minimise risk of misinterpretation by police officers. The current research investigated police officers' perceptions of the relations among contributory factors in the UK crash report form. While the current investigation focussed on the UK crash reporting practices as a case in point, the findings are intended to inform crash reporting practices in other countries by providing new insights into effective practices for developing road traffic collision reports.

2. Methods

2.1. Participants

One hundred sixty-two police officers (76% male; M_{age} = 42.85 years; age range = 24 − 64 years) were recruited from police stations and units across the United Kingdom. The author and representatives of the UK Department for Transport (DfT) contacted police stations and units across the UK. For police stations and units that agreed to take part, participation was solicited via email invitation sent to police officers via participating stations and units. Participants completed an online study of their views about road crash causation, lasting on average 24 min. Regarding their experience, officers indicated an average of 16.13 (SD = 8.12) years of experience attending or reporting on road traffic crashes and had attended or reported on an estimated 49 (SD = 110) road crashes in the past 12 months. The largest proportion (56%) indicated road policing as their specialist unit, followed by collision investigation (17%), and forensic collision investigation (5%). Regarding education, 21% indicated high school as their highest level of educational attainment, 45% indicated college or third level education (e.g., A-levels, diploma), 30% indicated an undergraduate degree, and 4% indicated a postgraduate degree. Ethical approval for the study protocol was provided by the internal ethics review board (institution: University of Essex; title: Assessment of the reliability of police reported road crash data; protocol number: JR1601). All participants provided written informed consent prior to participating in the study.

2.2. Materials and procedure

When a police officer reports on a road crash, they complete an crash report form that includes various crash details. In the crash report, police officers provide a subjective report of the factors that they believe contributed to the road crash. Officers can select among seven categories up to six contributory factors, including road environment (e.g., defective traffic signals), vehicle defects (e.g., defective breaks), injudicious actions (e.g., exceeding the speed limit), error or reaction (e.g., failured to look properly), impairment or distraction (e.g., driver using mobile phone), behaviour or inexperience (e.g., aggressive driving), and vision affected (e.g., dazzling headlights; UK DfT, 2018a). The crash reports are processed by local authorities (police, local authority, contractor) and are provided to the UK Department for Transport (DfT) for public use (UK DfT, 2018b).

Participants were told that the purpose of the research was to assess whether the existing categories in the current crash report form adequately reflect the factors they contain. They were provided a list of 63 factors without their associated category and were asked to categorise the factors according to how they believed the factors should be grouped. Participants assigned factors to categories by clicking and dragging the factors from the list with a mouse cursor into category boxes that appeared alongside the factor list. They could create up to 15 categories of related factors, but were asked to create only as many categories as they needed to assign all factors. Participants were explained that a category could include from one to as many factors as they wished. If they believed that a factor was unrelated to crash causation, they were instructed to leave it in its position without placing it in a category. Participants could also create their own factors if they believed that factors associated with crash causation were not covered by the 63 existing factors provided. To add a new factor, participants were required to create new factors from a list of blank editable factors.

3. Results

Police officers created a mean of 9.38 (SD = 2.62; min = 4, max = 15) categories, assigning a mean of 59.89 (SD = 8.62; min = 14, max = 63) of the 63 factors. Thus, officers typically believed that most (to all) of the 63 factors are relevant to road traffic crash causation and that the factors belong to multiple distinct categories.

To investigate police officers' judgements about the optimal category structure of the factors, their category assignments were submitted to an agglomerative hierarchical clustering analysis using the 'Cluster' package in R (Maechler et al., 2018). In agglomerative hierarchical clustering, each of the 63 contributory factors is first assigned to its own individual cluster. A dissimilarity matrix of the pairwise dissimilarities (i.e., distances) between each of the clusters is calculated and used to merge the clusters into increasingly inclusive clusters until all clusters are merged into a single cluster containing all 63 factors. Clusters that are least dissimilar (i.e., have the shortest distance) in the dissimilarity matrix are combined at each stage of clustering. Complete-linkage clustering using Gower distances in the dissimilarity matrix was adopted as this method is suited to categorical data (Maechler et al., 2018).

The optimal number of clusters was determined that maximised the internal coherence of the clusters (i.e., similarity of factors within clusters) and the external differences between clusters (i.e., dissimilarity of factors between clusters). Fig. 1 provides an Elbow plot of the sum of the squared distances within clusters as a measure of the dissimilarity between factors within clusters. As shown in Fig. 1, as the number of clusters increased, the sum of the squared differences within clusters decreased (i.e., similarity increased). The Elbow plot indicates an inflection point at seven clusters as increasing the number of clusters from seven (SS = 1.22) to eight (SS = 1.12) clusters showed a relatively small decrease in the sum of squared within-cluster distances, in comparison with six clusters (SS = 1.51). Thus, a seven-cluster solution provided an optimal number of clusters. However, the Elbow plot indicated a second inflection point at eleven clusters as increasing the number of clusters from eleven (SS = 0.77) to twelve (SS = 0.71) clusters showed a relatively small decrease in the sum of squared within-cluster distances, in comparison with ten clusters (SS = 0.90). Therefore, an eleven-cluster solution is also considered. Fig. 1 also provides a plot of the Silhouette coefficient, which is a measure of within-cluster consistency, with higher values indicating higher consistency. The Silhouette coefficient confirmed that within-cluster

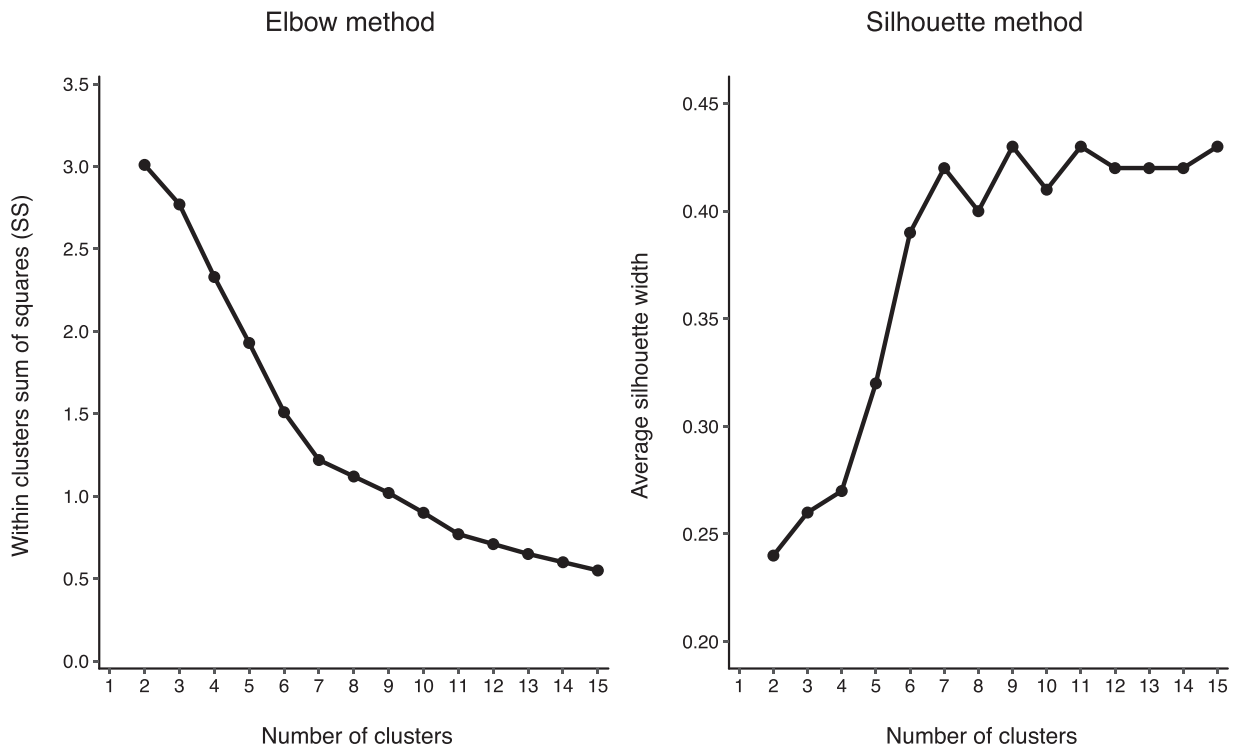

Fig. 1. Optimal number of clusters determined by the Elbow and Silhouette methods.

consistency was higher for a seven-cluster solution (coefficient = 0.42) compared to smaller numbers of clusters (Fig. 1), but showed a high coefficient also for an eleven-cluster solution (coefficient = 0.43).

3.1. Seven-cluster structure of contributory factors

Provided in Fig. 2 is a dendrogram showing the hierarchical structure of the seven clusters identified by our clustering analysis. The height of the links between factors within the dendrogram indicates the distance (i.e., dissimilarity) between factors in the dissimilarity matrix. For example, 'defective breaks' and 'defective steering or suspension' exhibited a small pairwise distance in the dissimilarity matrix, as indicated by the height of their link in the dendrogram (see Fig. 2), implying that they were often assigned by police officers to the same category. The hierarchical nature of the dendrogram shows how clusters begin as single-factor clusters and are successively combined into increasingly inclusive clusters until all factors are contained in a single cluster.

Observing Fig. 2, Cluster 1 contained five of the seven factors of the 'behaviour or inexperience' category in the crash report form (see Table 1). These five factors related to nervousness, uncertainty, and unfamiliarity, indicating that Cluster 1 represents an 'inexperience' category. This implies that police officers believe that a separate inexperience category captures road crash contributory factors that are distinct from other factors. The two remaining factors in the 'behaviour or inexperience' category in the crash report ('aggressive driving', 'careless, reckless, or in a hurry'), which relate to injudicious behaviour were instead clustered with the ten factors of the 'driver error or reaction' category and nine of the ten factors in the 'injudicious action' category in the crash report. Hence, Cluster 2 appears to represent an 'injudicious action or driver error' category, that includes related factors akin to poor or unwise judgement or behaviour that are distinct from other factors, such as those related to inexperience.

Cluster 3 contained all six factors of the 'vehicle defects' category in the crash report. The cluster also included one factor ('not displaying lights at night or in poor visibility') from the 'impairment or distraction' category, and one factor ('visor or windscreen dirty, scratched or frosted etc.') from the 'vision affected' category in the crash report. Thus, police officers believed that the latter two factors were more related to vehicle defects than to their original categories in the crash report, owing to their reference to defective features of the vehicle, namely vehicle lighting and visibility through the windscreen. Thus, Cluster 3 appears to represent a 'vehicle defects' category. Cluster 4 was comprised of eight of the ten factors of the 'impairment or distraction' category in the crash report, indicating that this cluster represents an 'impairment or distraction' category (Fig. 2). The remaining factor in the 'impairment or distraction' category of the crash report ('rider wearing dark clothing') was instead clustered with 'cyclist entering road from pavement' from the injudicious action category in Cluster 5, representing a 'cyclist error or visibility' category.

Cluster 6 contained nine of the ten factors from the 'vision affected' category in the crash report, implying that Cluster 6 is best described as a 'vision affected' category (Fig. 2). Cluster 7 contained all ten factors from the 'road Environment contributed' category in the crash report, and thus, is best described as a 'road environment contributed' category.

3.2. Eleven-cluster structure of contributory factors

Shown in Fig. 3 is a dendrogram identifying the eleven-cluster solution within the hierarchical structure identified in the clustering analysis. The hierarchical structure of the clusters is identical to the structure of the seven-factor solution, except that eleven, rather than seven, distinct clusters are identified. In the eleven-cluster structure, the 'injudicious action or driver error' category identified for the seven-factor structure, is further separated into a 'manoeuvring error' category, containing factors related to driver error during a manoeuvre ('sudden braking', 'swerved', 'loss or control', 'junction overshoot', 'junction restart [moving off at junction], 'poor turn or manoeuvre', and 'failed to signal or misleading signal'), a 'judgement error' category, containing factors related to poor judgement ('too close to cyclist, horse or pedestrian', 'failed to look properly', 'failed to judge other person's

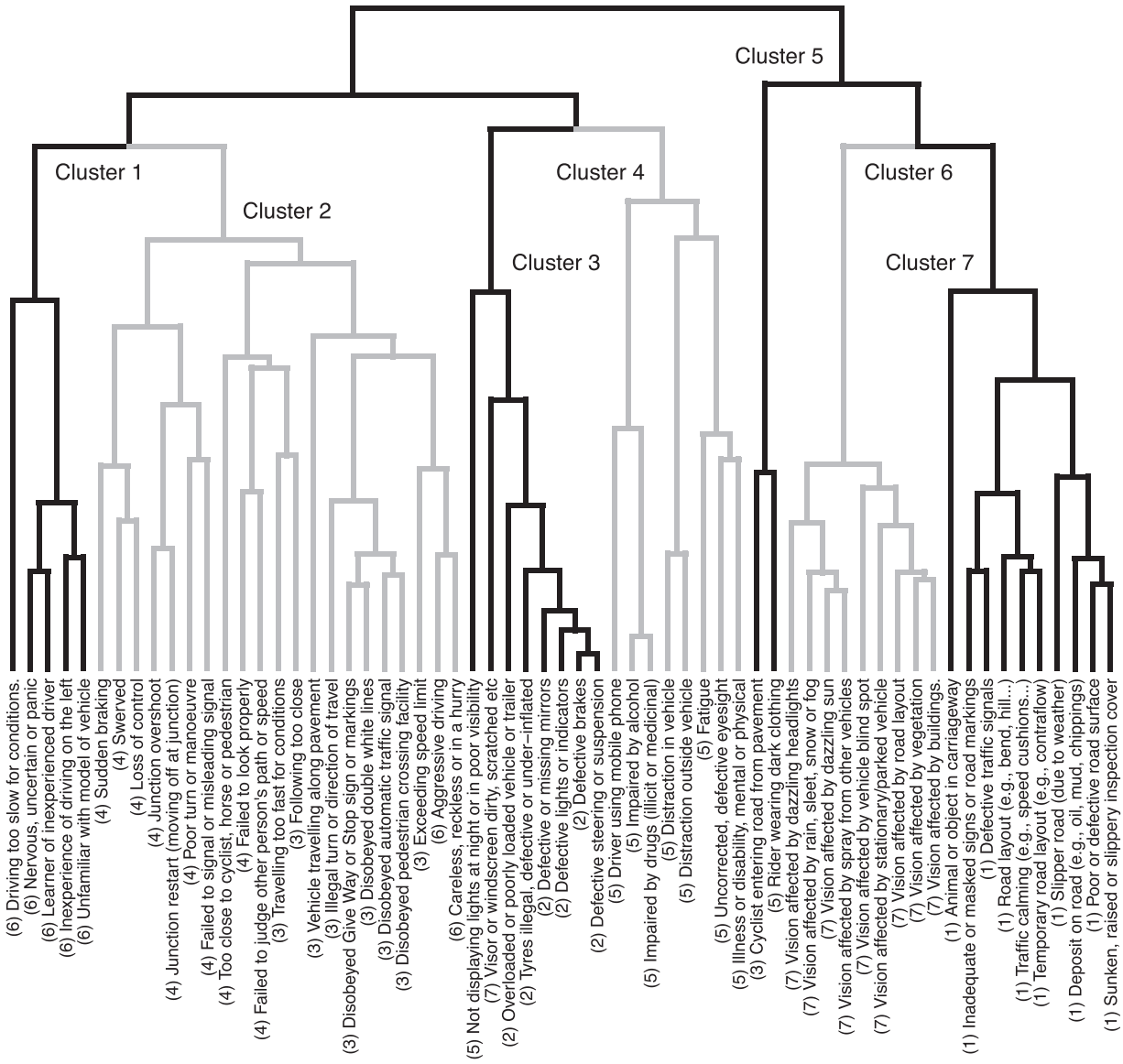

Fig. 2. Dendrogram showing the hierarchical structure of the seven-cluster solution containing all 63 contributory factors. In the contributory factor labels, 1 = 'road environment contributed', 2 = 'vehicle defects', 3 = 'injudicious action', 4 = 'driver error or reaction', 5 = 'impairment or distraction', 6 = 'behaviour or inexperience', and 7 = 'vision affected' in the UK road crash report form.

path or speed', 'traveling too fast for conditions', 'following too close'), and an 'injudicious action' category, containing factors related to poor judgement or unwise behaviour ('vehicle traveling along pavement', 'illegal turn or direction of travel', 'disobeyed give way or stop sign or markings', 'disobeyed double white lines', 'disobeyed automatic traffic signal', disobeyed pedestrian crossing facility', 'exceeding speed limit', 'aggressive driving', and 'careless, reckless, or in a hurry').

The 'impairment or distraction' category was further separated into a 'driver impairment' category, containing items related to impairment to driving ability ('driver using mobile phone', 'impairment by alcohol', impairment by drugs [illicit or medicinal]), a 'distraction' category, containing items related specifically to distracted driving ('distraction in vehicle', 'distraction outside vehicle'), and a 'driver fitness' category, containing factors related to the driver's physical fitness to drive ('fatigue', 'uncorrected, defective eyesight', 'illness or disability, mental or physical'; Fig. 3). Thus, the eleven-factor solution shows how two of the more eclectic categories can be separated into smaller, more coherent, categories.

3.3. Summary

Hierarchical clustering analysis on police officers' judgements about the categorical structure of contributory factors in the current UK road crash report form confirmed that most (to all) factors were considered relevant to crash causation. The optimal categorical structure of the factors broadly confirmed the structure employed in the existing crash report, but also revealed that some factors may be better placed in alternative categories. The hierarchical nature of the analytic approach revealed how larger categories (e.g., 'injudicious action or driver error') can be sub-divided into smaller categories.

4. Discussion

The current research investigated police officers' perceptions of the relations among contributory factors in a national road traffic collision report and employed hierarchical clustering analysis to identify the optimal category structure of the contributory factors. The clustering analysis identified a classification system with seven or eleven categories of contributory factors, maximising the internal coherence of

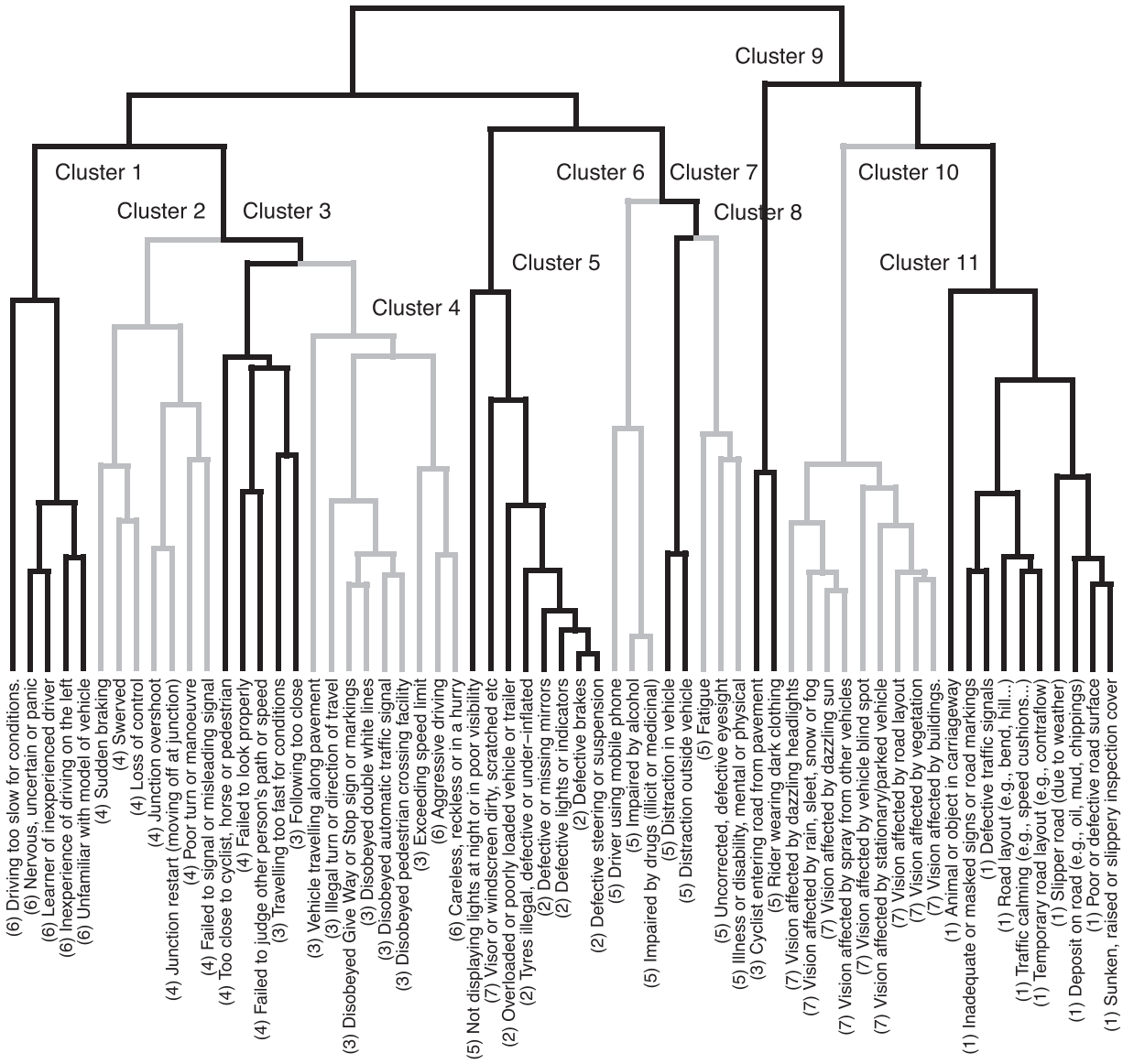

Fig. 3. Dendrogram showing the hierarchical structure of the eleven-cluster solution containing all 63 contributory factors. In the contributory factor labels, 1 = 'road environment contributed', 2 = 'vehicle defects', 3 = 'injudicious action', 4 = 'driver error or reaction', 5 = 'impairment or distraction', 6 = 'behaviour or inexperience', and 7 = 'vision affected' in the UK road crash report form.

categories (i.e., similarity among factors within categories). The findings yield new insights into police officers' perceptions of crash causation, as discussed below, and demonstrate how statistical clustering techniques can be used to inform the design of road crash report forms.

The current investigation focussed on the UK road traffic collision reporting procedures as a case in point with a view to informing the reporting of contributory factors internationally. In the UK crash report form, the 'behaviour or inexperience' category contains a variety of actions and behaviours, some of which refer explicitly to inexperience (e.g., 'learner or inexperienced driver') and others that do not refer to inexperience (e.g., 'careless, reckless, or in a hurry'; Table 1). The hierarchical clustering analysis, based on police officers' perceptions of the relations among the contributory factors, revealed a separate 'inexperience' category containing factors specifically related to driver inexperience (Cluster 1; Fig. 1). Hence, police officers perceived that contributory factors related to driver inexperience are qualitatively distinct from other driver actions or behaviours. Distinguishing inexperience-related factors in future developments of collision report forms should aid police officers in identifying appropriate contributory factors and reduce misinterpretations of the meaning of contributory factors. These findings have practical implications beyond the UK. In many countries, incident reports do not include a report of the contributory factors, unless a fatality occurs and an in-depth investigation is conducted. Police reports provide a rich data source for identifying patterns in crash causation. A practical implication of the current findings is that police officers' perceptions of the relations among contributory factors should be considered during the development of incident reports to provide the most accurate picture of crash causation.

The distinction between experience-related factors and other driver actions and behaviours also resonates with a focus in the road safety literature on young inexperienced drivers (Braitman et al., 2008; Rolison et al., 2013, 2014; Scott-Parker et al., 2012). The 'inexperience' category addresses a coherent set of behaviours related to a public health concern. The factors in the 'inexperience' category may even be suited to providing a naturalistic method of assessing the effectiveness of future education and training initiatives delivered regionally or nationally to target young driver behaviour, as methodological shortcomings have been identified in other methods of evaluation (Beanland et al., 2013b, 2013a;Rodwell et al., 2018). The 'learner or inexperienced driver' factor was strongly associated with the 'nervous,

uncertain, or panic' factor in police officers' perceptions, as indicated by the height of the link between these factors in the dendrogram (Fig. 2). Graduated licensing systems have been introduced in multiple countries, such as the United States, where young novice drivers are restricted to low-risk driving conditions (e.g., by restricting nighttime driving) to foster skill development (Bates et al., 2014; Shope, 2007). These systems could be assessed in terms of subsequent reduction in the occurrence of the 'learner or inexperienced driver' and 'nervous, uncertain, or panic' factors in road traffic collision reports. Further, the consistent use of experience-related factors in collision report forms internationally may also help foster national comparisons in road safety.

The UK crash report form does not distinguish a cyclist category from other contributory factor categories (Table 1). Rather, factors related to cyclists (e.g., 'cyclist entering road from pavement') are currently dispersed across categories (i.e., 'injudicious action'). Perceptions of police officers indicated a need for a separate cyclist category to distinguish cyclist-related factors from factors associated with other road users, increasing the coherence of the factor categories. Compared to car occupants, cyclists have a much higher risk of death or serious injury when involved in a road traffic collision (Lahrmann et al., 2018; Wegman et al., 2012). Cyclist visibility is an important determinant of cyclist risk as increasing visibility via bicycle lights or reflective clothing reduces risk of crash involvement (Lahrmann et al., 2018). Larger overall numbers of cyclists on the road also improves cyclist safety at least in part because cyclists are less likely to be overlooked by drivers (Fyhri et al., 2017). Hence, cyclist safety could be improved with practical interventions, such as increasing cyclist visibility and enhancing driver awareness. The current findings suggest that a separate contributory factor category dedicated to cyclists may help focus road safety priorities on cyclist safety and would provide a key outcome measure for assessing the impact of road safety policies and initiatives targeted at safeguarding cyclists. Cycling is a frequent mode of transport in many countries, such as in Denmark, Hungary, and the Netherlands, where it is the most frequent mode of transport for more than a fifth of people (European Commission, 2014). Cycling is also promoted in plans for sustainable city transport systems (e.g., European Commission, 2011). The current findings suggest that in the UK incident reporting practices, and in the development of such practices for the reporting of contributory factors in other countries, a separate contributory factor category dedicated to cyclists may help focus road safety priorities on cyclist safety. Doing so would also provide a key outcome measure for assessing the impact of road safety policies and initiatives targeted at safeguarding cyclists.

The hierarchical clustering identified cases where contributory factors were better placed in different categories to those used in the current UK crash reporting form. For example, the 'vehicle defects' category contained 'not displaying lights at night or in poor visibility', which currently appears in the 'impairment or distraction' category. This factor was not strongly associated with other factors more closely related to driver impairment (e.g., 'fatigue') or distraction (e.g., 'distraction in vehicle') according to police officers' perceptions, as indicated by the height of the links between these factors in the dendrogram (Fig. 2). The clustering analysis also indicated that the 'visor or windscreen dirty, scratched, or frosted etc.' factor, currently in the 'vision affected' category of the crash report form, is more appropriately placed in the 'vehicle defects' category. Indeed, this factor explicitly refers to vehicle defects. Psychological research has revealed that category membership alters perception of the items that categories contain (Goldstone and Hendrickson, 2010; Schusterman et al., 2000). When a category structure is imposed, such as categorising contributory factors, perceived differences between items in the same category are diminished and perceived differences between items in separate categories are exaggerated. Eclectic, or incoherent, contributory factor categories may minimise perceived differences between factors within a category, leading to misreporting due to misunderstanding of the meaning of individual contributory factors. Aligning the category structure with police officers' perceptions maximises category coherence, improving ease of crash reporting and reducing risk of reporting errors.

A positive feature of the hierarchical clustering approach employed in the current research is that clusters are successively combined into increasingly inclusive clusters. As such, clusters can be divided into smaller clusters without altering the overall structure (Maechler et al., 2018). The clustering analysis revealed an eleven-cluster solution, in which two categories in the seven-cluster solution were divided into less inclusive factor categories (Figures 2 & 3). In the eleven-cluster solution, the 'injudicious action or driver error' category was further divided into a 'manoeuvring error' category, a 'judgement error' category, and an 'injudicious action' category. Therefore, the 'injudicious action or driver error' category could be divided into smaller, more coherent, categories that focus on more specific aspects of driver actions and behaviour. The hierarchical structure also indicated that the 'impairment or distraction' category could be divided into an 'driver impairment' category and a 'driver fitness' category. Hence, the eleven-item category structure identifies more specific aspects of driver-related factors.

Intriguingly, in the eleven-cluster solution the 'driver using mobile phone' factor was clustered within the 'driver impairment' category with other factors related to impairment of the driver (e.g., 'impaired by alcohol'), rather than in the 'distraction' category with factors specifically related to distraction (e.g., 'distraction in vehicle'; Fig. 3). This suggests that police officers consider mobile phone use as more related to impairment of the driver than to distraction. Mobile phone use during driving is associated with impaired driving performance (Drews et al., 2008; Strayer et al., 2006). Even when using hands-free technology, mobile phone use impairs driver reactions by reducing attentional processing of the visual scene (Strayer et al., 2003). Hence, as police officers suggest, mobile phone use may be more accurately described as impairing driving ability, rather than distracting the driver from the act of driving, such as by averting their gaze. Mobile phone use is an under-reported factor in road crash records (NHTSA, 2009; Rolison et al., 2018). Categorising mobile phone use as a driver impairment, rather than a distraction, could help reduce under-reporting by better aligning its categorisation in the crash report form with police officers' perception of how it contributes to road crashes. On this basis, combined with a focus on more specific categories of factors, the eleven-item category structure may yield better insights into crash causation than the seven-item structure.

The current study has limitations. First, to identify the ideal category structure of the contributory factors, hierarchical clustering analysis was conducted on police officers' perceptions of the relations among the factors, rather than devise an objective method of estimating relations among the factors. In other words, police officers may be biased in their perceptions of contributory factors, leading to a biased category structure. However, the purpose of the current research was to devise a category structure that minimises discord with police officers' perceptions, in turn, reducing misinterpretation of the meaning of individual contributory factors. Hence, the proposed category structure provides an ergonomic design that is tailored to the user.

Second, police officers may have used their prior knowledge of the current UK road traffic collision reporting form, rather than their own perceptions of the contributory factors, to inform their grouping of the factors. Indeed, many features of the category structure devised from police officers' perceptions were compatible with the current collision reporting form. However, as discussed above, the perceptions of police officers also exhibited systematic differences from the current contributory factor category structure. These differences imply a discord between the current UK collision report form and the perceptions of police officers who use the form to report on road traffic collisions. While the category structure revealed by the current study shows advantages over the current report form in terms of its concordance with

police officers' perceptions, future research should seek to examine to what extent the new category structure improves incident reporting. A first step would be to assess whether collisions are reported differently with the new category structure compared to the current incident report form in terms of the factors selected by the police officer for a road traffic collision. By virtue of the closer alignment with police officers' perceptions of causation, the new category structure may also be easier and faster to complete by police officers than the current form, potentially leading to more efficient reporting and fewer reporting errors.

In conclusion, the current research investigated police officers' perceptions of the relations among contributory factors in the UK road traffic collision reporting form. Hierarchical clustering analysis revealed an optimal category structure of the factors that minimised discord with police officers' perceptions. The analysis also yielded new insights into police officers' perceptions of crash causation as well as demonstrating how statistical clustering analysis can be used to inform the design of road crash reports.

Author contributions

Jonathan Rolison designed the research, conducted the research and statistical analyses, and wrote the article.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research was supported by a grant awarded by the UK Engineering and Physical Sciences Research Council (EPSRC Reference; EP/M017877/1; "A new metric for the assessment of driver crash risks").

References

Bates, L.J., Allen, S., Armstrong, K., Watson, B., King, M.J., Davey, J., 2014. Graduated driver licensing: an international review. Sultan Qaboos Univ. Med. J. 14, e432.
Beanland, V., Fitzharris, M., Young, K.L., Lenné, M.G., 2013b. Driver inattention and driver distraction in serious casualty crashes: data from the Australian national crash in-depth study. Accid. Anal. Prev. 54, 99–107.
Beanland, V., Goode, N., Salmon, P.M., Lenné, M.G., 2013a. Is there a case for driver training? A review of the efficacy of pre-and post-licence driver training. Saf. Sci. 51, 127–137.
Braitman, K.A., Kirley, B.B., McCartt, A.T., Chaudhary, N.K., 2008. Crashes of novice teenage drivers: characteristics and contributing factors. J. Safety Res. 39, 47–54.
Broughton, J., Keigan, M., Knowles, J., Smith, L., 2010. Guidance for the Analysis of STATS19 Contributory Factors. Transportation Research Laboratory Available at: https://trl.co.uk/sites/default/files/PPR488.pdf. (Accessed 24 July, 2019).
Couto, A., Amorim, M., Ferreira, S., 2016. Reporting road victims: assessing and correcting data issues through distinct injury scales. J. Safety Res. 57, 39–45.
Drews, F.A., Pasupathi, M., Strayer, D.L., 2008. Passenger and cell phone conversations in simulated driving. J. Exp. Psychol. Appl. 14, 392–400.
Elder, R.W., Shults, R.A., Sleet, D.A., Nichols, J.L., Thompson, R.S., Rajab, W., Task Force on Community Preventive Services, 2004. Effectiveness of mass media campaigns for reducing drinking and driving and alcohol-involved crashes: a systematic review. Am. J. Prev. Med. 27, 57–65.
European Commission, 2011. European Commission (2011) WHITE PAPER Roadmap to a Single European Transport Area – Towards a Competitive and Resource Efficient Transport System. Available at: https://ec.europa.eu/transport/sites/transport/files/themes/strategies/doc/2011_white_paper/white-paper-illustrated-brochure_en.pdf. (Accessed 21 October, 2019). .
European Commission, 2014. Quality of Transport: Report. Available at: https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_422a_en.pdf. Accessed 21 October, 2019. .
Evans, L., 1996. The dominant role of driver behavior in traffic safety. Am. J. Public Health 86, 784–786.
Fyhri, A., Sundfør, H.B., Bjørnskau, T., Laureshyn, A., 2017. Safety in numbers for cyclists—conclusions from a multidisciplinary study of seasonal change in interplay and conflicts. Accid. Anal. Prev. 105, 124–133.
Goldstone, R.L., 1994. The role of similarity in categorization: providing a groundwork. Cognition 52, 125–157.
Goldstone, R.L., Hendrickson, A.T., 2010. Categorical perception. Wiley Interdiscip. Rev. Cogn. Sci. 1, 69–78.
Gonzales, M.M., Dickinson, L.M., DiGuiseppi, L.M., Lowenstein, S.R., 2005. Student drivers: a study of fatal motor vehicle crashes involving 16-year-old drivers. Ann. Emerg. Med. 45, 140–146.
Imprialou, M., Quddus, M., 2019. Crash data quality for road safety research: Current state and future directions. Accid. Anal. Prev. 130, 84–90.
Lam, L.T., 2003. Factors associated with young drivers' car crash injury: comparisons among learner, provisional, and full licensees. Accid. Anal. Prev. 35, 913–920.
Langford, J., Koppel, S., 2006. Epidemiology of older driver crashes–identifying older driver risk factors and exposure patterns. Transp. Res. Part F Traffic Psychol. Behav. 9, 309–321.
Larsen, L., 2004. Methods of multidisciplinary in-depth analyses of road traffic accidents. J. Hazard. Mater. 111, 115–122.
Lahrmann, H., Madsen, T.K.O., Olesen, A.V., 2018. Randomized trials and self-reported accidents as a method to study safety-enhancing measures for cyclists—two case studies. Accid. Anal. Prev. 114, 17–24.
Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., 2018. Cluster: cluster analysis basics and extensions. R package version 2 0.7–1.
McGwin, G., Brown, D.B., 1999. Characteristics of traffic crashes among young, middle-aged, and older drivers. Accid. Anal. Prev. 31, 181–198.
Montella, A., 2011. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. Accid. Anal. Prev. 43, 1451–1463.
NHTSA, 2008. National Motor Vehicle Crash Causation Survey: Report to Congress. Available at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811059. Accessed 18 October, 2019. .
NHTSA, 2009. Highlights of 2009 Motor Vehicle Crashes. Available at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811363. (Accessed 24 July, 2019). .
Petridou, E., Moustaki, M., 2000. Human factors in the causation of road traffic crashes. Eur. J. Epidemiol. 16, 819–826.
Rodwell, D., Hawkins, A., Haworth, N., Larue, G.S., Bates, L., Filtness, A., 2018. A mixed-methods study of driver education informed by the Goals for Driver Education: Do young drivers and educators agree on what was taught? Saf. Sci. 108, 140–148.
Rolison, J.J., Hewson, P.J., Hellier, E., Hurst, L., 2013. Risks of high-powered motorcycles among younger adults. Am. J. Public Health 103, 568–571.
Rolison, J.J., Moutari, S., Hewson, P.J., Hellier, E., 2014. Overestimated crash risks of young and elderly drivers. Am. J. Prev. Med. 46, 58–64.
Rolison, J.J., Regev, S., Moutari, S., Feeney, A., 2018. What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records. Accid. Anal. Prev. 115, 11–24.
Schusterman, R.J., Reichmuth, C.J., Kastak, D., 2000. How animals classify friends and foes. Curr. Dir. Psychol. Sci. 9, 1–6.
Scott-Parker, B., Watson, B., King, M.J., Hyde, M.K., 2012. Young, inexperienced, and on the road: Do novice drivers comply with road rules? Transp. Res. Rec. 2318, 98–106.
Shope, J.T., 2007. Graduated driver licensing: review of evaluation results since 2002. J. Safety Res. 38, 165–175.
Strayer, D.L., Drews, F.A., Crouch, D.J., 2006. A comparison of the cell phone driver and the drunk driver. Hum. Factors 48, 381–391.
Strayer, D.L., Drews, F.A., Johnston, W.A., 2003. Cell phone-induced failures of visual attention during simulated driving. J. Exp. Psychol. Appl. 9, 23–32.
UK Department for Transport (DfT, 2014. Reported Road Casualties Great Britain: 2014 Annual Report: Contributory Factors to Reported Road Accidents 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/463043/rrcgb2014-02.pdf. (Accessed 24 July, 2019). .
UK Department for Transport (DfT), 2017. Reported Road Casualties Great Britain: 2017 Annual Report. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755698/rrcgb-2017.pdf. Accessed 24 July, 2019. .
UK Department for Transport (DfT, 2018a. STATS19 Road Accident Injury Statistics: Report Form. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/230590/stats19.pdf. (Accessed 24 July, 2019). .
UK Department for Transport (DfT, 2018b. Instructions for the Completion of Road Accident Reports. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/230596/stats20-2011.pdf. (Accessed 24 July, 2019).
Watson, A., Watson, B., Vallmuur, K., 2015. Estimating under-reporting of road crash injuries to police using multiple linked data collections. Accid. Anal. Prev. 83, 18–25.
Wegman, F., Zhang, F., Dijkstra, A., 2012. How to make more cycling good for road safety? Accid. Anal. Prev. 44, 19–29.
Weijermars, W., Wegman, F., 2011. Ten years of sustainable safety in the Netherlands: an assessment. Trans. Res. Rec. 2213, 1–8.
World Health Organization, 2015. Global Status Report on Road Safety 2015. Available at: http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/. (Accessed July 29, 2019). .